12.31.97



OM protein - protein search, using sw model

Run on: April 8, 2003, 14:23:18 ; Search time 40 Seconds
(without alignments)
1815.540 Million cell updates/sec

Title: US-09-001-737-8
Perfect score: 2663
Sequence: 1 MAKEIKFSADARAAMVRGVD.....TPAPAMPAGMDPGMMGGMGG 545

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 2663 | 100.0 | 545 | 20 | AAY23904 | Streptococcus pyog |
| 2 | 2643 | 99.2 | 545 | 23 | ABP28529 | Streptococcus poly |
| 3 | 2405.5 | 90.3 | 540 | 23 | ABP28528 | Streptococcus poly |
| 4 | 2358 | 88.5 | 541 | 20 | AAY23902 | Streptococcus pneu |
| 5 | 2358 | 88.5 | 641 | 22 | AAB31619 | Amino acid sequenc |
| 6 | 2347 | 88.1 | 540 | 22 | AAM01101 | CFE 104 protein se |
| 7 | 2195.5 | 82.4 | 542 | 23 | ABB53701 | Lactococcus lactis |
| 8 | 2033.5 | 76.4 | 542 | 23 | ABB49241 | Listeria monocytog |
| 9 | 2026.5 | 76.1 | 544 | 20 | AAY23905 | Amino acid sequenc |
| 10 | 1904.5 | 71.5 | 545 | 20 | AAY23930 | Consensus mino aci |
| 11 | 1880.5 | 70.6 | 539 | 22 | AAG81848 | S. epidermidis ope |
| 12 | 1847.5 | 69.4 | 539 | 20 | AAY23916 | Amino acid sequenc |
| 13 | 1833.5 | 68.9 | 540 | 23 | ABP39860 | Staphylococcus epi |
| 14 | 1819 | 68.3 | 539 | 20 | AAY23906 | Amino acid sequenc |
| 15 | 1702 | 63.9 | 541 | 20 | AAY23910 | Amino acid sequenc |
| 16 | 1691.5 | 63.5 | 548 | 18 | AAW16678 | Lawsonia intracell |
| 17 | 1671 | 62.7 | 541 | 20 | AAY14909 | Amino acid sequenc |
| 18 | 1671 | 62.7 | 541 | 23 | ABB73515 | M vaccae GroEL hom |
| 19 | 1654 | 62.1 | 545 | 20 | AAY23915 | Amino acid sequenc |
| 20 | 1653.5 | 62.1 | 540 | 18 | AAW32100 | Mycobacteria sp. h |
| 21 | 1652.5 | 62.1 | 948 | 22 | AAB31611 | Amino acid sequenc |
| 22 | 1652 | 62.0 | 544 | 21 | AAY75747 | Neisseria meningit |
| 23 | 1647.5 | 61.9 | 540 | 9 | AAP81351 | Sequence of Mycoba |
| 24 | 1647.5 | 61.9 | 540 | 16 | AAR81610 | Mycobacterium tube |
| 25 | 1647.5 | 61.9 | 540 | 19 | AAW44702 | Mycobacterium tube |
| 26 | 1647.5 | 61.9 | 540 | 20 | AAY23911 | Amino acid sequenc |
| 27 | 1647.5 | 61.9 | 540 | 21 | AAY93332 | Amino acid sequenc |
| 28 | 1647.5 | 61.9 | 540 | 22 | AAE11755 | Mycobacterium tube |
| 29 | 1647.5 | 61.9 | 540 | 22 | AAG81118 | Mycobacterium tube |
| 30 | 1647.5 | 61.9 | 540 | 22 | AAB31606 | Amino acid sequenc |
| 31 | 1647.5 | 61.9 | 540 | 23 | AAM50750 | Mycobacterium tube |
| 32 | 1647.5 | 61.9 | 560 | 9 | AAP80215 | Sequence of Mycoba |
| 33 | 1647.5 | 61.9 | 639 | 22 | AAB31609 | Amino acid sequenc |
| 34 | 1647.5 | 61.9 | 648 | 22 | AAB31614 | Amino acid sequenc |
| 35 | 1643 | 61.7 | 523 | 19 | AAW60144 | M. vaccae antigen |
| 36 | 1643 | 61.7 | 523 | 20 | AAY14891 | Amino acid sequenc |
| 37 | 1643 | 61.7 | 523 | 23 | ABB73497 | M vaccae GroEL hom |
| 38 | 1642.5 | 61.7 | 638 | 21 | AAB03790 | Heat shock protein |
| 39 | 1641 | 61.6 | 545 | 20 | AAY23901 | Streptococcus pneu |
| 40 | 1640 | 61.6 | 548 | 19 | AAW65070 | E. coli GroEL prot |
| 41 | 1640 | 61.6 | 548 | 20 | AAW95001 | WO9902989 Seq ID 1 |
| 42 | 1640 | 61.6 | 548 | 22 | AAB50536 | Escherichia coli G |
| 43 | 1639.5 | 61.6 | 540 | 23 | AAU76511 | M. tuberculosis 65 |
| 44 | 1639.5 | 61.6 | 540 | 23 | AAU76194 | Mycobacterium tube |
| 45 | 1639.5 | 61.6 | 541 | 16 | AAR67384 | M. leprae GroEL1 g |

ALIGNMENTS

RESULT 1
AAY23904
ID AAY23904 standard; Protein; 545 AA.
XX
AC AAY23904;
XX
DT 22-SEP-1999 (first entry)
XX
DE Streptococcus pyogenes heat shock protein (Hsp)60-2.
XX
KW Heat shock protein; Hsp60-2; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO9935270-A1.
XX
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98WO-CA01203.
XX
PR 31-DEC-1997; 97US-0001737.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Mizzen L, Wisniewski J;
XX
DR WPI; 1999-430397/36.
DR N-PSDB; AAX86155.
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer
PT agents and for diagnosis

XX
PS Claim 11; Fig 4A-B; 176pp; English.
XX
CC The present sequence represents a heat shock protein, designated Hsp60-2.
CC The protein, its fragments, variants and fusion proteins, are
CC used to elicit or enhance an immune response against Streptococcus,
CC and to elicit a similar response to a target antigen fused to the
CC protein. Unlike other immunological carriers, Hsp60 proteins are not
CC immunosuppressive so provide an increased response to any conjugated or
CC fused antigen. Also, where used for cancer control, they lack the side
CC effects associated with endotoxins. They can also be used to detect
CC specific antibodies and in treatment or prevention of tumours
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC liver). The Hsp60 polynucleotide is used for recombinant production
CC of the protein, as a source of primers and probes for detecting
CC streptococci in standard hybridization/amplification assays, and
CC therapeutically in gene therapy vectors.
XX
SQ Sequence 545 AA;

Query Match 100.0%; Score 2663; DB 20; Length 545;
Best Local Similarity 100.0%; Pred. No. 6.9e-164;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60

Qy     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120

Qy    121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180

Qy    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240

Qy    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300

Qy    301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360

Qy    361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420

Qy    421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480

Qy    481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540

Qy    541 GGMGG 545
          |||||
Db    541 GGMGG 545
```

RESULT 2
ABP28529
ID ABP28529 standard; Protein; 545 AA.
XX
AC ABP28529;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 6234.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN69160.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3786; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 545 AA;

Query Match 99.2%; Score 2643; DB 23; Length 545;
Best Local Similarity 99.8%; Pred. No. 1.4e-162;
Matches 541; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
          |||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 MAKDIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 62

Qy     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 122

Qy    121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 182

Qy    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 242
```

```
Qy    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 302

Qy    301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 362

Qy    361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 422

Qy    421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 482

Qy    481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 542

Qy    541 GG 542
          ||
Db    543 GG 544
```

RESULT 3
ABP28528
ID  ABP28528 standard; Protein; 540 AA.
XX
AC  ABP28528;
XX
DT  02-JUL-2002 (first entry)
XX
DE  Streptococcus polypeptide SEQ ID NO 6232.
XX
KW  Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW  group A streptococcus; Streptococcus pyogenes; antibacterial;
KW  antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS  Streptococcus agalactiae.
XX
PN  WO200234771-A2.
XX
PD  02-MAY-2002.
XX
PF  29-OCT-2001; 2001WO-GB04789.
XX
PR  27-OCT-2000; 2000GB-0026333.
PR  24-NOV-2000; 2000GB-0028727.
PR  07-MAR-2001; 2001GB-0005640.
XX
PA  (CHIR-) CHIRON SPA.
PA  (GENO-) INST GENOMIC RES.
XX
PI  Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI  Tettelin H;
XX
DR  WPI; 2002-352536/38.
DR  N-PSDB; ABN69159.
XX
PT  New Streptococcus protein for the treatment or prevention of infection
PT  or disease caused by Streptococcus bacteria, such as meningitis, and
PT  for detecting a compound that binds to the protein -
XX
PS  Claim 1; Page 3785; 4525pp; English.
XX
CC  The invention relates to a protein (ABP25413-ABP30895) from group B
CC  streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC  (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC  the specification. The proteins have antibacterial and antiinflammatory
CC  activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC  antibodies that bind (I) are used in the manufacture of medicaments for
CC  the treatment or prevention of infection or disease caused by
CC  Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC  Nucleic acids encoding (I) are used to detect Streptococcus in a
CC  biological sample. (I) is used to determine whether a compound binds to
CC  (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC  used as a vaccine or diagnostic composition. The disease caused by
CC  Streptococcus that is prevented or treated may be meningitis. Nucleic
CC  acid encoding (I) may be used to recombinantly produce (I) and may be
CC  used in gene therapy. Antibodies to (I) are used for affinity
CC  chromatography, immunoassays, and distinguishing/identifying
CC  Streptococcus proteins.
XX
SQ  Sequence  540 AA;

Query Match  90.3%; Score 2405.5; DB 23; Length 540;
Best Local Similarity  90.2%; Pred. No. 2.9e-147;
Matches 489; Conservative 25; Mismatches 25; Indels 3; Gaps 2;

```
Qy      1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
          |||:||||||||:|||||||:|||||||||||||||||||||||||||||||||||||||
Db      1 MAKDIKFSADARSAMVRGVDILADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60

Qy     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVREGLKNVTAGANPIGIRRGIE 120

Qy    121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
          || : ||| || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TAVSAAVEELKEIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180

Qy    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
          ||||||||||||||||||||||||||||||::||||:||||||:||||:||||||||||
Db    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVSELENPYILITDKKISNIQEILPLLEEVLK 240

Qy    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVVT 300

Qy    301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
          |||||:|||||| |||:||:|||||||||||:| | |||||:|:||||:| |||||||
Db    301 EDLGLDLKDATMQVLGQSAKVTVDKDSTVIVEGAGDSSAIANRVAIIKSQMEATTSDFDR 360

Qy    361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
          |||||||||||||||||||| ||| ||||||||||||||||||||||:||||||: |
Db    361 EKLQERLAKLAGGVAVIKVGAATETELKEMKLRIEDALNATRAAVEEGIVSGGGTALVNV 420

Qy    421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
          |||||||:| ||: |||||||||||||||||| ||||||||:|::|| |  ||||||| |
Db    421 IEKVAALKLNGDEETGRNIVLRALEEPVRQIAYNAGYEGSVIIERLKQSEIGTGFNAANG 480

Qy    481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
          |||||: |||||||||||||||||||||||||||||||||||||||   ||  || ||| ||
Db    481 EWVDMVTTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEP--EAPTAPA-MDPSMM 537

Qy    541 GG 542
          ||
Db    538 GG 539
```

RESULT 4
AAY23902
ID  AAY23902 standard; Protein; 541 AA.
XX
AC  AAY23902;
XX
DT  22-SEP-1999 (first entry)
XX
DE  Streptococcus pneumoniae heat shock protein (Hsp)60-2.
XX
KW  Heat shock protein; Hsp60-2; immune response; immunological carrier;
KW  cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS  Streptococcus pneumoniae.
XX

PN WO9935270-A1.
XX
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98WO-CA01203.
XX
PR 31-DEC-1997; 97US-0001737.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Mizzen L, Wisniewski J;
XX
DR WPI; 1999-430397/36.
DR N-PSDB; AAX86153.
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer
PT agents and for diagnosis
XX
PS Claim 11; Fig 2A-B; 176pp; English.
XX
CC The present sequence represents a heat shock protein, designated Hsp60-2.
CC The protein, its fragments, variants and fusion proteins, are
CC used to elicit or enhance an immune response against Streptococcus,
CC and to elicit a similar response to a target antigen fused to the
CC protein. Unlike other immunological carriers, Hsp60 proteins are not
CC immunosuppressive so provide an increased response to any conjugated or
CC fused antigen. Also, where used for cancer control, they lack the side
CC effects associated with endotoxins. They can also be used to detect
CC specific antibodies and in treatment or prevention of tumours
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
CC liver). The Hsp60 polynucleotide is used for recombinant production
CC of the protein, as a source of primers and probes for detecting
CC streptococci in standard hybridization/amplification assays, and
CC therapeutically in gene therapy vectors.
XX
SQ Sequence 541 AA;

Query Match 88.5%; Score 2358; DB 20; Length 541;
Best Local Similarity 87.9%; Pred. No. 3.3e-144;
Matches 479; Conservative 33; Mismatches 29; Indels 4; Gaps 1;

```
Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |||||||:|||:|||||||:|||||||||||||||||:|||||||||||||||||||
Db       1 MAKEIKFSSDARSAMVRGVDILADTVKVTLGPKGRNVVLEKSFGSPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           ||||||||||||||||||||||||||||||||||||||| ||:||||||||||||||||||
Db      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVREGIKNVTAGANPIGIRRGIE 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
           || | ||||||    ||: ||||||||||||||||||||||||:|| ||||||||||||
Db     121 TAVAAAVEALKNNVIPVANKEAIAQVAAVSSRSEKVGEYISEAMEKVGKDGVITIEESRG 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
           ||||||||||||||||||||||:||||||||||:|||||||:||||:|||||| :|:
Db     181 METELEVVEGMQFDRGYLSQYMVTDSEKMVADLENPYILITDKKISNIQEILPLLESILQ 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
           :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 SNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
           ||||||||||: |||||::|||||||||||||:|: |||::|:|:||||:||||||:|||
Db     301 EDLGLELKDATIEALGQAARVTVDKDSTVIVEGAGNPEAISHRVAVIKSQIETTTSEFDR 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
           ||||||||||:|||||||| ||| ||||||||||||||||||||||||||||||  |
Db     361 EKLQERLAKLSGGVAVIKVGAATETELKEMKLRIEDALNATRAAVEEGIVAGGGTALANV 420

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
           |  || ||| ||:|||||||||||||||||| |||:|||:|||:|||:  | |||||||
Db     421 IPAVATLELTGDEATGRNIVLRALEEPVRQIAHNAGFEGSIVIDRLKNAELGIGFNAATG 480

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
           |||:||  ||||||||:|||||||||||||||||||||||||||  ||||    ||| ||
Db     481 EWVNMIDQGIIDPVKVSRSALQNAASVASLILTTEAVVANKPEPVAPAPA----MDPSMM 536

Qy     541 GGMGG 545
           |||||
Db     537 GGMGG 541
```

RESULT 5
AAB31619
ID AAB31619 standard; Protein; 641 AA.
XX
AC AAB31619;
XX
DT 30-APR-2001 (first entry)
XX
DE Amino acid sequence of Hsp65-E7 fusion protein.
XX
KW Heat shock protein; Hsp; Th1 response; Th1 cell; CD4+ T lymphocyte cell;
KW lymphocyte; Hsp65; Hsp40; Hsp10; Hsp60; Hsp71; microbial pathogen;
KW E7 protein.
XX
OS Synthetic.
OS Streptococcus pneumoniae.
OS Human papillomavirus.
XX
PN WO200104344-A2.
XX
PD 18-JAN-2001.
XX
PF 10-JUL-2000; 2000WO-US18828.
XX
PR 08-JUL-1999; 99US-0143757.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Siegel M, Chu NR, Mizzen LA;
XX
DR WPI; 2001-138361/14.
DR N-PSDB; AAF25036.
XX
PT Screening for compounds that stimulate Th1-like responses in CD4+ T
PT lymphocyte cells
XX
PS Example 15; Fig 15A-B; 88pp; English.
XX
CC The present sequence represents a fusion protein comprising a
CC Streptococcus pneumoniae heat shock protein (Hsp) 65 fused to a HPV16 E7
CC protein. The fusion protein is used in the method of the invention. The
CC specification describes a method of determining whether a compound
CC stimulates a Th1-like response. Th1 cells are a subset of CD4+
CC T lymphocyte cells. The method comprises contacting naive lymphocytes
CC in vitro with a fusion protein comprising at least a fragment of Hsp,
CC and then detecting the TH1-like response exhibited by the cell sample.
CC The proteins which may be used in the method of the invention are Hsp65,
CC Hsp40, Hsp10, Hsp60, and Hsp71. The method may be used to identify
CC compounds that stimulate Th1-like responses in response to microbial
CC pathogens.
XX
SQ Sequence 641 AA;

Query Match 88.5%; Score 2358; DB 22; Length 641;
Best Local Similarity 87.9%; Pred. No. 4.1e-144;
Matches 479; Conservative 33; Mismatches 29; Indels 4; Gaps 1;

```
Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |||||||:|||:|||||||:|||||||||||||||||:|||||||||||||||||||
Db       1 MAKEIKFSSDARSAMVRGVDILADTVKVTLGPKGRNVVLEKSFGSPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           ||||||||||||||||||||||||||||||||||||||| ||:||||||||||||||||||
```

```
Db     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVREGIKNVTAGANPIGIRRGIE 120

Qy    121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
          || | ||||||    ||: ||||||||||||||||||||||||||:|| ||||||||||
Db    121 TAVAAAVEALKNNVIPVANKEAIAQVAAVSSRSEKVGEYISEAMEKVGKDGVITIEESRG 180

Qy    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
          |||||||||||||||||||||||||:||||||||||:|||||||:||||:|||||| :|:
Db    181 METELEVVEGMQFDRGYLSQYMVTDSEKMVADLENPYILITDKKISNIQEILPLLESILQ 240

Qy    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300

Qy    301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
          ||||||||||: |||||::|||||||||||||:|: |||::|:|:||||:|||||:|||
Db    301 EDLGLELKDATIEALGQAARVTVDKDSTVIVEGAGNPEAISHRVAVIKSQIETTTSEFDR 360

Qy    361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
          ||||||||||:|||||||| ||| ||||||||||||||||||||||||||||  |
Db    361 EKLQERLAKLSGGVAVIKVGAATETELKEMKLRIEDALNATRAAVEEGIVAGGGTALANV 420

Qy    421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
          |  || ||| ||:||||||||||||||||||| |||:|||:|||:|||:  | |||||||
Db    421 IPAVATLELTGDEATGRNIVLRALEEPVRQIAHNAGFEGSIVIDRLKNAELGIGFNAATG 480

Qy    481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
          |||:||  |||||||:|||||||||||||||||||||||||  |||| |||  ||
Db    481 EWVNMIDQGIIDPVKVSRSALQNAASVASLILTTEAVVANKPEPVAPAPA----MDPSMM 536

Qy    541 GGMGG 545
          |||||
Db    537 GGMGG 541


RESULT 6
AAM01101
ID   AAM01101 standard; Protein; 540 AA.
XX
AC   AAM01101;
XX
DT   02-OCT-2001  (first entry)
XX
DE   CFE 104 protein sequence.
XX
KW   Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KW   CFE; CEG; Conserved Essential Gene; bacterial infection;
KW   antisense therapy; antibiotic resistance.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO200149721-A2.
XX
PD   12-JUL-2001.
XX
PF   29-DEC-2000; 2000WO-US35604.
XX
PR   30-DEC-1999;   99US-0174089.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Dougherty TJ,  Pucci MJ,  Dougherty BA,  Davison DB,  Bruccoleri RE;
PI   Thanassi JA;
XX
DR   WPI; 2001-496721/54.
DR   N-PSDB; AAH90800.
XX
PT   Nucleic acids encoding conserved essential genes involved in bacterial
PT   replication which are potential targets for the treatment of antibiotic
PT   resistant bacterial infections -
XX
PS   Claim 27; Pages 356-358; 380pp; English.
XX
CC   The present invention relates to nucleic acids (AAH90701-AAH90918)
CC   encoding polypeptides (AAM01002-AAM01114), which are essential for the
CC   viability of a bacterial cell wall. The acronym CFE stands for "CEG For
CC   Expression", where CEG stands for "Conserved Essential Gene". The nucleic
CC   acids are useful for detecting the presence of proteins essential for the
CC   viability of a bacterial cell wall in samples such as cells, tissues,
CC   biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC   and for detecting corresponding target nucleic acid molecules with
CC   complementary sequences. The nucleic acids are also useful for
CC   determining whether a genomic nucleotide sequence of interest is
CC   essential for viability of a bacterial cell or whether it resides within
CC   an operon, by integrating an exogeneous nucleotide sequence comprising a
CC   portion of an open reading frame of the genomic sequence of interest
CC   (comprising 200-500 base pairs) into the genomic sequence of interest
CC   which confers a selectable phenotype to the cell, and determining cell
CC   viability with a selection agent such as chloramphenicol. The nucleic
CC   acids and proteins are also useful as vaccines and for treating bacterial
CC   infections with gene therapy and antisense therapy. The nucleic acids
CC   also enable identification of targets suitable for the treatment of
CC   antibiotic resistant bacterial infections.
XX
SQ   Sequence   540 AA;

  Query Match             88.1%;  Score 2347;  DB 22;  Length 540;
  Best Local Similarity   87.8%;  Pred. No. 1.7e-143;
  Matches  477;  Conservative   34;  Mismatches   28;  Indels    4;  Gaps    1;

Qy      1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
          |:|||||:|||:|||||||:||||||||||||||||||||:||||||||||||||||||
Db      1 MSKEIKFSSDARSAMVRGVDILADTVKVTLGPKGRNVVLEKSFGSPLITNDGVTIAKEIE 60

Qy     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
          |||||||||||||||||||||||||||||||||||||| ||:||||||||||||||||||
Db     61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVREGIKNVTAGANPIGIRRGIE 120

Qy    121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
          || | ||||||  | ||: ||||||||||||||||||||||||:|| ||||||||||
Db    121 TAVAAAVEALKNNAIPVANKEAIAQVAAVSSRSEKVGEYISEAMEKVGKDGVITIEESRG 180

Qy    181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
          |||||||||||||||||||||||||||:||||||||||:|||||||:||||:|||||| :|:
Db    181 METELEVVEGMQFDRGYLSQYMVTDSEKMVADLENPYILITDKKISNIQEILPLLESILQ 240

Qy    241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
          :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 SNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300

Qy    301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
          ||||||||||: |||||::|||||||||||||:|: |||::|:|:||||:|||||:|||
Db    301 EDLGLELKDATIEALGQAARVTVDKDSTVIVEGAGNPEAISHRVAVIKSQIETTTSEFDR 360

Qy    361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
          ||||||||||:|||||||| ||| ||||||||||||||||||||||||||||  |
Db    361 EKLQERLAKLSGGVAVIKVGAATETELKEMKLRIEDALNATRAAVEEGIVAGGGTALANV 420

Qy    421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
          |  || ||| ||:||||||||||||||||||| |||:|||:|||:|||:  | |||||||
Db    421 IPAVATLELTGDEATGRNIVLRALEEPVRQIAHNAGFEGSIVIDRLKNAELGIGFNAATG 480

Qy    481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
          |||:||  |||||||:|||||||||||||||||||||||||  |||| |||  ||
Db    481 EWVNMIDQGIIDPVKVSRSALQNAASVASLILTTEAVVANKPEPVAPAPA----MDPSMM 536

Qy    541 GGM 543
          |||
Db    537 GGM 539


RESULT 7
ABB53701
ID   ABB53701 standard; Protein; 542 AA.
XX
AC   ABB53701;
```

```
XX
DT   16-MAY-2002  (first entry)
XX
DE   Lactococcus lactis protein groEL.
XX
KW   Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
OS   Lactococcus lactis IL1403.
XX
PN   FR2807446-A1.
XX
PD   12-OCT-2001.
XX
PF   11-APR-2000; 2000FR-0004630.
XX
PR   11-APR-2000; 2000FR-0004630.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Bolotine A,  Sorokine A,  Renault P,  Ehrlich SD;
XX
DR   WPI; 2002-043418/06.
XX
PT   New nucleotide sequence useful in the identification or Lactococcus
PT   lactis and related species -
XX
PS   Claim 6; SEQ ID No 403; 2504pp; French.
XX
CC   The present invention is related to a Lactococcus lactis nucleotide
CC   sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC   nucleic acid sequence is useful in the detection and/or amplification of
CC   nucleic acid sequence, particularly to identify Lactococcus lactis or
CC   related species. The proteins of the invention are useful for the
CC   biosynthesis or biodegradation of a composition of interest. The
CC   invention helps research in lactic bacteria, particularly useful in the
CC   production of yogurt and cheese.
CC   Note: The sequence data for this patent is based on equivalent patent
CC   WO200177334 (published 18-OCT-2001) which is available in electronic
CC   format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   542 AA;

  Query Match             82.4%;  Score 2195.5;  DB 23;  Length 542;
  Best Local Similarity   81.1%;  Pred. No. 1e-133;
  Matches  442;  Conservative   52;  Mismatches   46;  Indels    5;  Gaps    3;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |:|:||||:||| ||:||:|:|||||| |||||||||||||::|||||||||||||||||
Db       1 MSKDIKFSSDARTAMMRGIDILADTVKTTLGPKGRNVVLEKSYGSPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||:|||||||
Db      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVREGLKNVTAGANPVGIRRGIE 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
            |  ||| ::| :| ||  | |||||| ||||||||||||:||||||:||||||||:|
Db     121 LAAETAVASIKEMAIPVHDKSAIAQVATVSSRSEKVGEYISDAMERVGSDGVITIEESKG 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
           |:|||:|||||||||||||||:: |||||:|:||:|||||||:||||:|||||::||
Db     181 MQTELDVVEGMQFDRGYLSQYMVSNTEKMVAELDNPYILITDKKISNIQEILPLLEQILK 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
           |||||||:|||||||||||||||:| |||||||||||||||| |||:||||||||||
Db     241 TNRPLLIVADDVDGEALPTLVLNKIKGVFNVVAVKAPGFGDRRKAQLEDLAILTGGTVIT 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
           |:|||:|||||: ||||||| ||||| | ||||:||::||::|:|:||:|:| |||||||
Db     301 EELGLDLKDATLEALGQAAKATVDKDHTTIVEGAGSADAISDRVAIIKAQIEKTTSDFDR 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
           |||||||||||||||:|||| ||| || ||| ||||||||||||||||:||||||:
Db     361 EKLQERLAKLAGGVAVVKVGAATETELKAMKLLIEDALNATRAAVEEGIVSGGGTALVNA 420

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
           |  :  |  |||  || ||| |||||||||| ||||||||:||||::   |||||||||
Db     421 IAALDKLSEEGDIQTGINIVRRALEEPVRQIAANAGYEGSVIIDKLRSEEVGTGFNAATG 480

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
           :||:||: ||:|| |||||||||||||| ||||||||||||||||  |||||  ||| |
Db     481 QWVNMIEEGIVDPAKVTRSALQNAASVAGLILTTEAVVANKPEPA--APAMPP-MDPSM- 536

Qy     541 GGMGG 545
             ||||
Db     537 -GMGG 540


RESULT 8
ABB49241                                                              1
ID   ABB49241 standard; Protein; 542 AA.
XX
AC   ABB49241;
XX
DT   05-FEB-2002  (first entry)
XX
DE   Listeria monocytogenes protein #1945.
XX
KW   Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW   vitamin B12; bacterial infection; disease.
XX
OS   Listeria monocytogenes.
XX
PN   WO200177335-A2.
XX
PD   18-OCT-2001.
XX
PF   11-APR-2001; 2001WO-FR01118.
XX
PR   11-APR-2000; 2000FR-0004629.
XX
PA   (INSP ) INST PASTEUR.
XX
PI   Buchrieser C,  Frangeul L,  Couve E,  Rusniok C,  Fsihi H,  Dehoux P;
PI   Dussurget O,  Chetouani F,  Nedjari H,  Glaser P,  Kunst F,  Cossart P;
PI   Daniels J,  Goebel W,  Kreft J,  Kuhn M,  Ng E,  Vazquez-Boland JA;
PI   Dominguez-Bernal G,  Garrido-Garcia P,  Tierrez-Martinez A,  Amend A;
PI   Chakraborty T,  Domann E,  Hain T,  Berche P,  Charbit A,  Durant L;
PI   Perez-Diaz J,  Baquero F,  Garcia Del Portillo F,  Gomez-Lopez N;
PI   Maduenio E,  De Pablos B,  Wehland J,  Kaerst U,  Entian K,  Hauf J;
PI   Rose M,  Voss H;
XX
DR   WPI; 2002-010914/01.
XX
PT   Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT   and prevention of Listeria and related bacterial infections, and
PT   related polypeptides  -
XX
PS   Claim 6; SEQ ID No 1946; 192pp; French.
XX
CC   The present invention relates to the genome sequence of Listeria
CC   monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC   it are useful for selecting probes and primers for detecting genes in L.
CC   monocytogenes and related organisms, and for studying genetic
CC   polymorphisms and other genomes. The present sequence is a protein
CC   encoded by the genome sequence of the present invention. Proteins
CC   expressed from the genome sequence are useful for raising specific
CC   antibodies, identification of L. monocytogenes and related organisms, and
CC   for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC   B12. The genome sequence and proteins encoded by it are also useful for
CC   selecting compounds that regulate gene expression and cell replication
CC   and modulate L. monocytogenes-related diseases. In addition, the genome
CC   sequence and proteins encoded by it are useful in pharmaceutical and
CC   vaccines compositions for the treatment or prevention of infections by L.
CC   monocytogenes and related organisms.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
```

```
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   542 AA;

  Query Match           76.4%;  Score 2033.5;  DB 23;  Length 542;
  Best Local Similarity  75.6%;  Pred. No. 2.9e-123;
  Matches  412;  Conservative   61;  Mismatches   67;  Indels    5;  Gaps    3;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |||:|||| ||| ||:||:|| ||: |||||||||||||||| ||||||||||||||||||
Db       1 MAKDIKFSEDARRAMLRGVDQLANAVKVTLGPKGRNVVLEKKFGSPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           ||| |||||||||||||||||:|||||||||||| ||:: |||||||||||:|:|||||
Db      61 LEDPFENMGAKLVSEVASKTNDVAGDGTTTATVLAQAMIQEGLKNVTAGANPVGVRRGIE 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
             | |||:| ||||:::|:  ||:|||||:|| |:||: |:||||||||||||||||:|
Db     121 KAVATAIEELKAISKPIESKESIAQVAAISSGDEEVGKLIAEAMERVGNDGVITIEESKG 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
           |||:|||||||||| | |||||::|| | || |:|||||||::|||:|||:||:|::
Db     181 FATELDVVEGMQFDRGYTSPYMVTDSDKMEAVLEKPYILITDKKINNIQEILPVLEQVVQ 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
             ||:||||:||:|||  |||||:|||||||||||||||||||||||||||||| |||
Db     241 QGRPMLIIAEDVEGEAQATLVLNKLRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGQVIT 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
           |||||||| ||:  || | |: | || | |||||:| |  |: |:  |::|:| |||:|||
Db     301 EDLGLELKTATVDQLGTANKVVVTKDDTTIVEGAGDSTQISARVNQIRAQMEETTSEFDR 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
           |||||||||||||||:|||| ||| ||| |||||||||:|||||||||||||||||:::
Db     361 EKLQERLAKLAGGVAVVKVGAATETELKERKLRIEDALNSTRAAVEEGIVAGGGTALVSI 420

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
           |||||| ||| ||||||:|||||||| ||| ||||:::||:   | ||||| |
Db     421 YNKVAALEAEGDVETGINIVLRSLEEPVRQIAHNAGLEGSVIVERLKHEAVGVGFNAANG 480

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
           |||:|| ||:|| ||||||||||:|||:|:|||||||:||:   || |:|   | | |
Db     481 EWVNMIDAGIVDPTKVTRSALQNASSVAALLLTTEAVVADKPDENGPA-AVP---DMG-M 535

Qy     541 GGMGG 545
           |||||
Db     536 GGMGG 540


RESULT 9
AAY23905
ID   AAY23905 standard; Protein; 544 AA.
XX
AC   AAY23905;
XX
DT   22-SEP-1999  (first entry)
XX
DE   Amino acid sequence of a heat shock protein.
XX
KW   Heat shock protein; Hsp; immune response; immunological carrier;
KW   cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS   Bacillus subtilis.
XX
PN   WO9935270-A1.
XX
PD   15-JUL-1999.
XX
PF   29-DEC-1998;   98WO-CA01203.
XX
PR   31-DEC-1997;   97US-0001737.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Wisniewski J;
XX
DR   WPI; 1999-430397/36.
XX
PT   New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT   useful in vaccines, as carriers for other immunogens, as anticancer
PT   agents and for diagnosis
XX
PS   Disclosure; Fig 10A-E; 176pp; English.
XX
CC   AAY23905-30 represent heat shock proteins (Hsps). The specification
CC   describes Streptococcal Hsps, designated Hsp60. These proteins, their
CC   fragments, variants and fusion proteins, are used to elicit or enhance
CC   an immune response against Streptococcus, and to elicit a similar
CC   response to a target antigen fused to the protein. Unlike other
CC   immunological carriers, Hsp60 proteins are not immunosuppressive so
CC   provide an increased response to any conjugated or fused antigen. Also,
CC   where used for cancer control, they lack the side effects associated
CC   with endotoxins. They can also be used to detect specific antibodies
CC   and in treatment or prevention of tumours (e.g. sarcoma or cancers of
CC   breast, ovary, prostate, lung, pancreas or liver). The Hsp60
CC   polynucleotide is used for recombinant production of the protein, as
CC   a source of primers and probes for detecting streptococci in standard
CC   hybridization/amplification assays, and therapeutically in gene
CC   therapy vectors.
XX
SQ   Sequence   544 AA;

  Query Match           76.1%;  Score 2026.5;  DB 20;  Length 544;
  Best Local Similarity  73.8%;  Pred. No. 8.3e-123;
  Matches  402;  Conservative   69;  Mismatches   71;  Indels    3;  Gaps    1;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |||||||| :|| ||:|||| ||| ||||||||||||||| ||||||||||||||||||
Db       1 MAKEIKFSEEARRAMLRGVDALADAVKVTLGPKGRNVVLEKKFGSPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           ||| |||||||||:|||||||:|||||||||| ||:: |||||||||||:|:|:|:|
Db      61 LEDAFENMGAKLVAEVASKTNDVAGDGTTTATVLAQAMIREGLKNVTAGANPVGVRKGME 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
             | | |:| || |::|: |||:|||||:|:  |:||  |:|||||||||||||||:|
Db     121 QAVAVAIENLKEISKPIEGKESIAQVAAISAADEEVGSLIAEAMERVGNDGVITIEESKG 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
           |||||||||||||||| | |||||::|| | |:||:|||||||::|||:|||:||:|::
Db     181 FTTELEVVEGMQFDRGYASPYMVTDSDKMEAVLDNPYILITDKKITNIQEILPVLEQVVQ 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
             :|||:||:||:|||| |||:||:||||| ||||||||||||||||:|||| |||
Db     241 QGKPLLLIAEDVEGEALATLVVNKLRGTFNAVAVKAPGFGDRRKAMLEDIAVLTGGEVIT 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
           |||||:||   :  ||:|:|: | |::| ||||:| :: |: |:  |::|:| |||:|||
Db     301 EDLGLDLKSTQIAQLGRASKVVVTKENTTIVEGAGETDKISARVTQIRAQVEETTSEFDR 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
           |||||||||||||||||||| ||| ||| |||||||||:|||||||||:|||||: |
Db     361 EKLQERLAKLAGGVAVIKVGAATETELKERKLRIEDALNSTRAAVEEGIVSGGGTALVNV 420

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
           ||||:| |||  || ||||||||||:|||| ||| ||||::::|||   | |||||||
Db     421 YNKVAAVEAEGDAQTGINIVLRALEEPIRQIAHNAGLEGSVIVERLKNEEIGVGFNAATG 480

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
           |||:||: ||:|| ||||||||||||:: ||||||||:|||       ||   | | |
Db     481 EWVNMIEKGIVDPTKVTRSALQNAASVAAMFLTTEAVVADKPEENGGGAGMP---DMGGM 537

Qy     541 GGMGG 545
           |||||
Db     538 GGMGG 542
```

RESULT 10
AAY23930
ID AAY23930 standard; Protein; 545 AA.
XX
AC AAY23930;
XX
DT 22-SEP-1999 (first entry)
XX
DE Consensus mino acid sequence of a heat shock protein.
XX
KW Heat shock protein; Hsp; immune response; immunological carrier;
KW cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS Synthetic.
XX
PN WO9935270-A1.
XX
PD 15-JUL-1999.
XX
PF 29-DEC-1998; 98WO-CA01203.
XX
PR 31-DEC-1997; 97US-0001737.
XX
PA (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI Mizzen L, Wisniewski J;
XX
DR WPI; 1999-430397/36.
XX
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT useful in vaccines, as carriers for other immunogens, as anticancer
PT agents and for diagnosis
XX
PS Disclosure; Fig 10A-E; 176pp; English.
XX
CC AAY23905-30 represent heat shock proteins (Hsps). The specification
CC describes Streptococcal Hsps, designated Hsp60. These proteins, their
CC fragments, variants and fusion proteins, are used to elicit or enhance
CC an immune response against Streptococcus, and to elicit a similar
CC response to a target antigen fused to the protein. Unlike other
CC immunological carriers, Hsp60 proteins are not immunosuppressive so
CC provide an increased response to any conjugated or fused antigen. Also,
CC where used for cancer control, they lack the side effects associated
CC with endotoxins. They can also be used to detect specific antibodies
CC and in treatment or prevention of tumours (e.g. sarcoma or cancers of
CC breast, ovary, prostate, lung, pancreas or liver). The Hsp60
CC polynucleotide is used for recombinant production of the protein, as
CC a source of primers and probes for detecting streptococci in standard
CC hybridization/amplification assays, and therapeutically in gene
CC therapy vectors.
XX
SQ Sequence 545 AA;

Query Match 71.5%; Score 1904.5; DB 20; Length 545;
Best Local Similarity 71.7%; Pred. No. 6.3e-115;
Matches 392; Conservative 66; Mismatches 80; Indels 9; Gaps 6;

```
Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |||:|||  :|| ||:|||: ||| |||||||||||||||:||:| || ||||:|||||
Db       1 MAKDIKFGEEARRAMLRGVNALADAVKVTLGPKGRNVVLEKSFGAPTITKDGVTVAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           ||| ||||||:|| |||||||||:|||||||||| |||| |||||| ||||: :|||||:
Db      61 LEDKFENMGAQLVKEVASKTNDVAGDGTTTATVLAQAIVKEGLKNVAAGANPMDLRRGID 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSR-SEKVGEYISEAMERVGNDGVITIEESR 179
             |    || |||||:||  || ||||| :|:   |::|| |:||||:|| :||||:|| :
Db     121 KAVDAVVEELKAIAKPVETKEEIAQVATISANGDEEIGELIAEAMEKVGKEGVITVEEGK 180

Qy     180 GMETELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVL 239
           :|||||||||||||||:| | :||:||  |:||:| ||:|||||:||||||:||:|||||
Db     181 TLETELEVVEGMQFDRGYISPYFITDSEKQKAELEDPLILLTDKKISNIQDLLPVLEEVA 240

Qy     240 KTNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVI 299
           :  :|||||||:||:|||| |||:||:|||  |||||||||||||||||:|||||||| ||
Db     241 QAGKPLLIIAEDVEGEALATLVVNKLRGTLKVVAVKAPGFGDRRKAMLQDIAILTGGQVI 300

Qy     300 TEDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFD 359
           :|:||| |:|||:  |||| |: | || | ||:|:|: : ||| |:| |:||:| :|||:|
Db     301 SEELGLSLEDATLEDLGQAKKVVVTKDDTTIVDGAGDA-AIAGRVAQIRSQIEESTSDYD 359

Qy     360 REKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALIT 419
           :|||||||||||||||||||| ||  ||| | |:|||||||||||||| ||| ||:
Db     360 KEKLQERLAKLAGGVAVIKVGAATEVELKERKDRVEDALNATRAAVEEGIVPGGGVALLR 419

Qy     420 VIEKVAALELE-GDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAA 478
               :  |: | ||:||| |||||||| |:|||| ||| |||||::|:||| || |:|||
Db     420 AAPALDKLKTENGDEATGVNIVLRALEAPLRQIAENAGLEGSVVVEKVKNSEAG-GYNAA 478

Qy     479 TGEWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPG 538
           |||:|||| |||| |||||||||||||||:||||||| :|||      | |||| ||
Db     479 TGEYVDMIAAGIIDPTKVTRSALQNAASVASLMLTTEAVVVDKPE----KEAAPAGM-PG 533

Qy     539 MMGGMGG 545
           |||||||
Db     534 MMGGMGG 540
```

RESULT 11
AAG81848
ID AAG81848 standard; Protein; 539 AA.
XX
AC AAG81848;
XX
DT 03-SEP-2001 (first entry)
XX
DE S. epidermidis open reading frame protein sequence SEQ ID NO:790.
XX
KW Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW vaccination; endocarditis.
XX
OS Staphylococcus epidermidis.
XX
PN WO200134809-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30782.
XX
PR 09-NOV-1999; 99US-0164258.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Kimmerly WJ;
XX
DR WPI; 2001-316495/33.
DR N-PSDB; AAH52698.
XX
PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 18; Page 243; 2188pp; English.
XX
CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA

```
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence   539 AA;

  Query Match             70.6%;  Score 1880.5;  DB 22;  Length 539;
  Best Local Similarity   68.1%;  Pred. No. 2.2e-113;
  Matches  368;  Conservative   85;  Mismatches   86;  Indels    1;  Gaps    1;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
Db       1 MAKDLKFSEDARQAMLRGVDKLANAVKVTIGPKGRNVVLDKDYTTPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
Db      61 LEDPYENMGAKLVQEVANKTNEIAGDGTTTATVLAQSMIQEGLKNVTSGANPVGLRQGID 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
Db     121 KAVQVAIEALHEISQKVENKNEIAQVGAISAADEEIGRYISEAMDKVGNDGVITIEESNG 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
Db     181 FNTELEVVEGMQFDRGYQSPYMVTDSDKMIAELERPYILVTDKKISSFQDILPLLEQVVQ 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
Db     241 ASRPILIVADEVEGDALTNIVLNRMRGTFTAVAVKAPGFGDRRKAMLEDLAILTGAQVIT 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
Db     301 DDLGLELKDASLDMLGTANKVEVTKDHTTVVDGNGDENNIDARVGQIKAQIEETDSEFDK 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
Db     361 EKLQERLAKLAGGVAVIKVGAASETELKERKLRIEDALNSTRAAVEEGIVAGGGTALVNI 420

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
Db     421 YQKVSEIKAEGDVETGVNIVLKALQAPVRQIAENAGLEGSIIVERLKHAEAGVGFNAATN 480

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
Db     481 EWVNMLEEGIVDPTKVTRSALQHAASVAAMFLTTEAVVASIPEPENNEQPGMGGM-PGMM 539


RESULT 12
AAY23916
ID   AAY23916 standard; Protein; 539 AA.
XX
AC   AAY23916;
XX
DT   22-SEP-1999  (first entry)
XX
DE   Amino acid sequence of a heat shock protein.
XX
KW   Heat shock protein; Hsp; immune response; immunological carrier;
KW   cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS   Staphylococcus aureus.
XX
PN   WO9935270-A1.
XX
PD   15-JUL-1999.
XX
PF   29-DEC-1998;   98WO-CA01203.
XX
PR   31-DEC-1997;   97US-0001737.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Wisniewski J;
XX
DR   WPI; 1999-430397/36.
XX
PT   New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT   useful in vaccines, as carriers for other immunogens, as anticancer
PT   agents and for diagnosis
XX
PS   Disclosure; Fig 10A-E; 176pp; English.
XX
CC   AAY23905-30 represent heat shock proteins (Hsps). The specification
CC   describes Streptococcal Hsps, designated Hsp60. These proteins, their
CC   fragments, variants and fusion proteins, are used to elicit or enhance
CC   an immune response against Streptococcus, and to elicit a similar
CC   response to a target antigen fused to the protein. Unlike other
CC   immunological carriers, Hsp60 proteins are not immunosuppressive so
CC   provide an increased response to any conjugated or fused antigen. Also,
CC   where used for cancer control, they lack the side effects associated
CC   with endotoxins. They can also be used to detect specific antibodies
CC   and in treatment or prevention of tumours (e.g. sarcoma or cancers of
CC   breast, ovary, prostate, lung, pancreas or liver). The Hsp60
CC   polynucleotide is used for recombinant production of the protein, as
CC   a source of primers and probes for detecting streptococci in standard
CC   hybridization/amplification assays, and therapeutically in gene
CC   therapy vectors.
XX
SQ   Sequence   539 AA;

  Query Match             69.4%;  Score 1847.5;  DB 20;  Length 539;
  Best Local Similarity   68.3%;  Pred. No. 2.9e-111;
  Matches  371;  Conservative   78;  Mismatches   87;  Indels    7;  Gaps    4;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
Db       1 MVKQLKFSEDARQAMLRGVDQLANAVKVTIGPKGRNVVLDKEFTAPLITNDGVTIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
Db      61 LEDPYENMGAKLVQEVANKTNEIAGDGTTTATVLAQAMIQEGLKNVTSGANPVGLRQGID 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGV---ITIEE 177
Db     121 KAVKVAVEALHENSQKVENKNEIAQVGAISAADEEIGRYISEATEKVGNDGVITIITIEE 180

Qy     178 SRGMETELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEE 237
Db     181 SNRLNTELEL--GMQFDRGYQSPYMVTDSDKMVAELERPYILVTDKKISSFQDILPLLEQ 238

Qy     238 VLKTNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGT 297
Db     239 VVQSNRPILIVADEVEGDALTNIVLNRMRGTFTAVAVKAPGFGDRRKAMLEDLAILTGAQ 298

Qy     298 VITEDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSD 357
Db     299 VITDDLGLDLKDASIDMLGTASKVEVTKDNTTVVDGDGDENSIDARVSQLKSQIEETESD 358

Qy     358 FDREKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTAL 417
Db     359 FDREKLQERLAKLAGGVAVIKVGAASETELKERKLRIEDALNSTRAAVEEGIVAGGGTAL 418

Qy     418 ITVIEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNA 477
Db     419 VNVYQKVSENEAEGDIETGVNIVLKALTAPVRQIAENAGLEGSVIVERLKNAEPGVGFNG 478

Qy     478 ATGEWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDP 537
Db     479 ATNEWVNMLRRGIVDPTKVTRSALQHAASVAAMFLTTEAVVASIPEKNNDQPNM-GGM-P 536

Qy     538 GMM 540
```

```
Db    537 GMM 539


RESULT 13
ABP39860
ID   ABP39860 standard; Protein; 540 AA.
XX
AC   ABP39860;
XX
DT   24-JUL-2002  (first entry)
XX
DE   Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4705.
XX
KW   Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
KW   antibacterial; gene therapy.
XX
OS   Staphylococcus epidermidis.
XX
PN   US6380370-B1.
XX
PD   30-APR-2002.
XX
PF   13-AUG-1998;   98US-0134001.
XX
PR   14-AUG-1997;   97US-055779P.
PR   08-NOV-1997;   97US-064964P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm LA,  Bush D;
XX
DR   WPI; 2002-381255/41.
DR   N-PSDB; ABN92405.
XX
PT   Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT   polypeptide, useful for diagnosing and treating bacterial infections -
XX
PS   Disclosure; SEQ ID 4705; 267pp; English.
XX
CC   ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC   frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC   given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC   antibacterial activity and can be used in gene therapy. The sequences
CC   can also be used in the diagnosis and treatment of bacterial infections,
CC   particularly S. epidermidis infections. The sequences can be used to
CC   screen for compounds able to interfere with the S. epidermidis life
CC   cycle or inhibit S. epidermidis infection.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from the
CC   USPTO web site.
XX
SQ   Sequence   540 AA;

  Query Match             68.9%;  Score 1833.5;  DB 23;  Length 540;
  Best Local Similarity   66.5%;  Pred. No. 2.4e-110;
  Matches  359;  Conservative   87;  Mismatches   93;  Indels    1;  Gaps    1;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
           |||::|||| ||| ||:|||| ||: ||||:||||||||:| : :|||||||||||||||
Db       2 MAKDLKFSEDARQAMLRGVDKLANAVKVTIGPKGRNVVLDKDYTTPLITNDGVTIAKEIE 61

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
           ||| :|||||||| |||:|||:|||||||||||| |::: |||||||:||||:|:|:||:
Db      62 LEDPYENMGAKLVQEVANKTNEIAGDGTTTATVLAQSMIQEGLKNVTSGANPVGLRQGID 121

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
           |   |:||| |:| |  |  |||| |:|:  |::| |||||::||||||||||| |
Db     122 KAVQVAIEALHEISQKVENKNEIAQVGAISAADEEIGRYISEAMDKVGNDGVITIEESNG 181

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
           |||||||||||||| | |||||::||:|:|| |:|||:||||:|: ||||||||:|::
Db     182 FNTELEVVEGMQFDRGYQSPYMVTDSDKMIAELERPYILVTDKKISSFQDILPLLEQVVQ 241

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
           :|||:||:||:|:|:||  :|||::|||| ||||||||||||||||:|||||  |||
Db     242 ASRPILIVADEVEGDALTNIVLNRMRGTFTAVAVKAPGFGDRRKAMLEDLAILTGAQVIT 301

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
           :|||||||||::  || | |: | || | :|:|:|    |  |:  ||:|:| | |:||:
Db     302 DDLGLELKDASLDMLGTANKVEVTKDHTTVVDGNGDENNIDARVGQIKAQIEETDSEFDK 361

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
           ||: | | || ||||| :|   :|| ||| ||||||||:|||||||||||||||||: :
Db     362 EKITESLGKLPGGVAVNQVWGGSETELKERKLRIEDALNSTRAAVEEGIVAGGGTALVNI 421

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
           :|||: :: ||| || ||||:||: |||||| ||| |||:::::||:: || ||||||
Db     422 YQKVSEIKAEGDVETGVNIVLKALQAPVRQIAENAGLEGSIIVERLKHAEAGVGFNAATN 481

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
           |||:|:: ||:|| |||||||:|||||:: |||||||: ||| || ||||
Db     482 EWVNMLEEGIVDPTKVTRSALQHAASVAAMFLTTEAVVASIPEPENNEQPGMGGM-PGMM 540


RESULT 14
AAY23906
ID   AAY23906 standard; Protein; 539 AA.
XX
AC   AAY23906;
XX
DT   22-SEP-1999  (first entry)
XX
DE   Amino acid sequence of a heat shock protein.
XX
KW   Heat shock protein; Hsp; immune response; immunological carrier;
KW   cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS   Clostridium sp.
XX
PN   WO9935270-A1.
XX
PD   15-JUL-1999.
XX
PF   29-DEC-1998;   98WO-CA01203.
XX
PR   31-DEC-1997;   97US-0001737.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Wisniewski J;
XX
DR   WPI; 1999-430397/36.
XX
PT   New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT   useful in vaccines, as carriers for other immunogens, as anticancer
PT   agents and for diagnosis
XX
PS   Disclosure; Fig 10A-E; 176pp; English.
XX
CC   AAY23905-30 represent heat shock proteins (Hsps). The specification
CC   describes Streptococcal Hsps, designated Hsp60. These proteins, their
CC   fragments, variants and fusion proteins, are used to elicit or enhance
CC   an immune response against Streptococcus, and to elicit a similar
CC   response to a target antigen fused to the protein. Unlike other
CC   immunological carriers, Hsp60 proteins are not immunosuppressive so
CC   provide an increased response to any conjugated or fused antigen. Also,
CC   where used for cancer control, they lack the side effects associated
CC   with endotoxins. They can also be used to detect specific antibodies
CC   and in treatment or prevention of tumours (e.g. sarcoma or cancers of
CC   breast, ovary, prostate, lung, pancreas or liver). The Hsp60
CC   polynucleotide is used for recombinant production of the protein, as
CC   a source of primers and probes for detecting streptococci in standard
CC   hybridization/amplification assays, and therapeutically in gene
CC   therapy vectors.
XX
```

```
SQ   Sequence   539 AA;

  Query Match           68.3%;  Score 1819;  DB 20;  Length 539;
  Best Local Similarity   67.7%;  Pred. No. 2e-109;
  Matches  367;  Conservative   77;  Mismatches   92;  Indels    6;  Gaps    2;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
Db       1 MAKTLLFGEEARRSMQAGVDKLANTVKVTLGPKGRNVILDKKFGSPLITNDGVTIAREIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
Db      61 LEDAYENMGAQLVKEVATKTNDVAGDGTTTATLLAQAIIREGLKNVTAGANPILIRNGIK 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
Db     121 TAVEKAVEEIQKISKPVNGKEDIARVAAISAADEKIGKLIADAMEKVGNEGVITVEESKS 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
Db     181 MGTELDVVEGMQFDRGYVSAYMVTDTEKMEAVLDNPLVLITDKKISNIQDLLPLLEQIVQ 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
Db     241 AGKKLLIIADDIEGEAMTTLVVNKLRGTFTCVGVKAPGFGDRRKEMLQDIATLTGGVVIS 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
Db     301 DEVGGDLKEATLDMLGEAESVKVTKESTTIVNGRGNSEEIKNRINQIKLQLEATTSEFDK 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
Db     361 EKLQERLAKLAGGVAVVKVGAATETELKESKLRIEDALAATKAAVEEGIVPGGGTAYVNV 420

Qy     421 IEKVAALELE-GDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAAT 479
Db     421 INEVAKLTSDIQDEQVGINIIVRSLEEPMRQIAHNAGLEGSVIIEKVKNSDAGVGFDALR 480

Qy     480 GEWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGM 539
Db     481 GEYKDMIKAGIVDPTKVTRSALQNAASVASTFLTTEAAVADIPE-----KEMPQGAGMGM 535

Qy     540 MG 541
Db     536 DG 537


RESULT 15
AAY23910
ID   AAY23910 standard; Protein; 541 AA.
XX
AC   AAY23910;
XX
DT   22-SEP-1999  (first entry)
XX
DE   Amino acid sequence of a heat shock protein.
XX
KW   Heat shock protein; Hsp; immune response; immunological carrier;
KW   cancer control; tumour; sarcoma; cancer; gene therapy.
XX
OS   Mycobacterium avium.
XX
PN   WO9935270-A1.
XX
PD   15-JUL-1999.
XX
PF   29-DEC-1998;   98WO-CA01203.
XX
PR   31-DEC-1997;   97US-0001737.
XX
PA   (STRE-) STRESSGEN BIOTECHNOLOGIES CORP.
XX
PI   Mizzen L,  Wisniewski J;
XX
DR   WPI; 1999-430397/36.
XX
PT   New nucleic acid encoding heat shock protein-60 from Streptococcus,
PT   useful in vaccines, as carriers for other immunogens, as anticancer
PT   agents and for diagnosis
XX
PS   Disclosure; Fig 10A-E; 176pp; English.                1
XX
CC   AAY23905-30 represent heat shock proteins (Hsps). The specification
CC   describes Streptococcal Hsps, designated Hsp60. These proteins, their
CC   fragments, variants and fusion proteins, are used to elicit or enhance
CC   an immune response against Streptococcus, and to elicit a similar
CC   response to a target antigen fused to the protein. Unlike other
CC   immunological carriers, Hsp60 proteins are not immunosuppressive so
CC   provide an increased response to any conjugated or fused antigen. Also,
CC   where used for cancer control, they lack the side effects associated
CC   with endotoxins. They can also be used to detect specific antibodies
CC   and in treatment or prevention of tumours (e.g. sarcoma or cancers of
CC   breast, ovary, prostate, lung, pancreas or liver). The Hsp60
CC   polynucleotide is used for recombinant production of the protein, as
CC   a source of primers and probes for detecting streptococci in standard
CC   hybridization/amplification assays, and therapeutically in gene
CC   therapy vectors.
XX
SQ   Sequence   541 AA;

  Query Match           63.9%;  Score 1702;  DB 20;  Length 541;
  Best Local Similarity   63.0%;  Pred. No. 7.3e-102;
  Matches  342;  Conservative   84;  Mismatches  113;  Indels    4;  Gaps    1;

Qy       1 MAKEIKFSADARAAMVRGVDMLADTVKVTLGPKGRNVVLEKAFGSPLITNDGVTIAKEIE 60
Db       1 MAKTIAYDEEARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEIE 60

Qy      61 LEDHFENMGAKLVSEVASKTNDIAGDGTTTATVLTQAIVHEGLKNVTAGANPIGIRRGIE 120
Db      61 LEDPYEKIGAELVKEVAKKTDDVAGDGTTTATVLAQALVREGLRNVAAGANPLGLKRGIE 120

Qy     121 TATATAVEALKAIAQPVSGKEAIAQVAAVSSRSEKVGEYISEAMERVGNDGVITIEESRG 180
Db     121 KAVEKVTETLLKSAKEVETKDQIAATAAISAGDQSIGDLIAEAMDKVGNEGVITVEESNT 180

Qy     181 METELEVVEGMQFDRGYLSQYMVTDNEKMVADLENPFILITDKKVSNIQDILPLLEEVLK 240
Db     181 FGLQLELTEGMRFDKGYISGYFVTDAERQEAVLEDPFILLVSSKVSTVKDLLPLLEKVIQ 240

Qy     241 TNRPLLIIADDVDGEALPTLVLNKIRGTFNVVAVKAPGFGDRRKAMLEDIAILTGGTVIT 300
Db     241 AGKPLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGGQVIS 300

Qy     301 EDLGLELKDATMTALGQAAKITVDKDSTVIVEGSGSSEAIANRIALIKSQLETTTSDFDR 360
Db     301 EEVGLSLESADISLLGKARKVVVTKDETTIVEGAGDSDAIAGRVAQIRTEIENSDSDYDR 360

Qy     361 EKLQERLAKLAGGVAVIKVGAPTETALKEMKLRIEDALNATRAAVEEGIVAGGGTALITV 420
Db     361 EKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVALLHA 420

Qy     421 IEKVAALELEGDDATGRNIVLRALEEPVRQIALNAGYEGSVVIDKLKNSPAGTGFNAATG 480
Db     421 IPALDELKPEGEEATGANIVRVALERPLKQIAFNGGLEPGVVAEKVRNSPAGTGLNAATG 480

Qy     481 EWVDMIKTGIIDPVKVTRSALQNAASVASLILTTEAVVANKPEPATPAPAMPAGMDPGMM 540
Db     481 EYEDLLKAGIADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKA----AAPAGDPTGGM 536

Qy     541 GGM 543
Db     537 GGM 539


Search completed: April  8, 2003, 14:24:29
```

Job time : 43 secs